

Quantitative limit theorems for local functionals of arithmetic random waves

Giovanni Peccati and Maurizia Rossi

Abstract We consider Gaussian Laplace eigenfunctions on the two-dimensional flat torus (arithmetic random waves), and provide explicit Berry-Esseen bounds in the 1-Wasserstein distance for the normal and non-normal high-energy approximation of the associated Leray measures and total nodal lengths, respectively. Our results provide substantial extensions (as well as alternative proofs) of findings by Oravecz, Rudnick and Wigman (2007), Krishnapur, Kurlberg and Wigman (2013), and Marinucci, Peccati, Rossi and Wigman (2016). Our techniques involve Wiener-Itô chaos expansions, integration by parts, as well as some novel estimates on residual terms arising in the chaotic decomposition of geometric quantities that can implicitly be expressed in terms of the coarea formula.

1 Introduction

The high-energy analysis of local geometric quantities associated with the *nodal set* of random Laplace eigenfunctions on compact manifolds has gained enormous momentum in recent years, in particular for its connections with challenging open problems in differential geometry (such as *Yau's conjecture* [19]), and with the striking cancellation phenomena detected by Berry in [2] — see the survey [18] for an overview of this domain of research up to the year 2012, and the Introduction of [13] for a review of recent literature. The aim of this paper is to prove quantitative limit theorems, in the high-energy limit, for *nodal lengths* and *Leray measures*

Giovanni Peccati
e-mail: giovanni.peccati@gmail.com

Maurizia Rossi
e-mail: maurizia.rossi@uni.lu

Unité de Recherche en Mathématiques, Université du Luxembourg
6, rue Coudenhove-Kalergi, L-1359 Luxembourg

(analogous to occupation densities at zero) of Gaussian Laplace eigenfunctions on the two-dimensional flat torus. These random fields, first introduced by Rudnick and Wigman in [16], are called *arithmetic random waves* and are the main object discussed in the paper. The term ‘arithmetic’ emphasises the fact that, in the two dimensional case, the definition of toral eigenfunctions is inextricable from the problem of enumerating lattice points lying on circles with integer square radius.

Our results will allow us, in particular, to recover by an alternative (and mostly self-contained) approach the variance estimates from [11], as well as the non-central limit theorems proved in [13]. The core of our approach relies on the use of the Malliavin calculus techniques described in the monograph [14], as well as on some novel combinatorial estimates for residual terms arising in variance estimates obtained by chaotic expansions.

Although the analysis developed in the present paper focusses on a specific geometric model, we reckon that our techniques might be suitably modified in order to deal with more general geometric objects, whose definitions involve some variation of the area/coarea formulae; for instance, we believe that one could follow a route similar to the one traced below in order to deduce quantitative versions of the non-central limit theorems for phase singularities proved in [5], as well as to recover the estimates on the nodal variance of toral eigenfunctions and random spherical harmonics, respectively deduced in [16] and [17].

From now on, every random object is supposed to be defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with \mathbb{E} denoting expectation with respect to \mathbb{P} .

1.1 Setup

As anticipated, in this paper we are interested in proving quantitative limit theorems for geometric quantities associated with Gaussian eigenfunctions of the Laplace operator $\Delta := \partial^2/\partial x_1^2 + \partial^2/\partial x_2^2$ on the flat torus $\mathbb{T} := \mathbb{R}^2/\mathbb{Z}^2$. In order to introduce our setup, we start by defining

$$S := \{n \in \mathbb{Z} : n = a^2 + b^2, \text{ for some } a, b \in \mathbb{Z}\}$$

to be the set of all numbers that can be written as a sum of two integer squares. It is a standard fact that the eigenvalues of $-\Delta$ are of the form $4\pi^2 n =: E_n$, where $n \in S$. The dimension \mathcal{N}_n of the eigenspace \mathcal{E}_n corresponding to the eigenvalue E_n coincides with the number $r_2(n)$ of ways in which n can be expressed as the sum of two integer squares (taking into account the order of summation). The quantity $\mathcal{N}_n = r_2(n)$ is a classical object in arithmetics, and is subject to large and erratic fluctuations: for instance, it grows *on average* as $\sqrt{\log n}$ but could be as small as 8 for an infinite sequence of prime numbers $p_n \equiv 1 \pmod{4}$, or as large as a power of $\log n$ – see [10, Section 16.9 and Section 16.10] for a classical discussion, as well as [12] for recent advances. We also set

$$\Lambda_n := \{\lambda = (\lambda_1, \lambda_2) \in \mathbb{Z}^2 : |\lambda|^2 := \lambda_1^2 + \lambda_2^2 = n\}$$

to be the class of all lattice points on the circle of radius \sqrt{n} (its cardinality $|\Lambda_n|$ equals \mathcal{N}_n). Note that Λ_n is invariant w.r.t. rotations around the origin by $k \cdot \pi/2$, where k is any integer. An orthonormal basis $\{e_\lambda\}_{\lambda \in \Lambda_n}$ for the eigenspace \mathcal{E}_n is given by the complex exponentials

$$e_\lambda(x) := \exp(i2\pi\langle\lambda, x\rangle), \quad x = (x_1, x_2) \in \mathbb{T}.$$

We now consider a collection (indexed by the set of frequencies $\lambda \in \Lambda_n$) of identically distributed standard complex Gaussian random variables $\{a_\lambda\}_{\lambda \in \Lambda_n}$, that we assume to be independent except for the relations $\overline{a_\lambda} = a_{-\lambda}$. We recall that, by definition, every a_λ has the form $a_\lambda = b_\lambda + ic_\lambda$, where b_λ, c_λ are i.i.d. real Gaussian random variables with mean zero and variance $1/2$. We define the *arithmetic random wave* [11, 13, 15] of order $n \in S$ to be the real-valued centered Gaussian function

$$T_n(x) := \frac{1}{\sqrt{\mathcal{N}_n}} \sum_{\lambda \in \Lambda_n} a_\lambda e_\lambda(x), \quad x \in \mathbb{T}; \quad (1)$$

from (1) it is easily checked that the covariance of T_n is given by, for $x, y \in \mathbb{T}$,

$$r_n(x, y) := \mathbb{E}[T_n(x) \cdot T_n(y)] = \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} \cos(2\pi\langle\lambda, x - y\rangle) =: r_n(x - y). \quad (2)$$

Note that $r_n(0) = 1$, i.e. $T_n(x)$ has unit variance for every $x \in \mathbb{T}$. Moreover, as emphasised in the right-hand side (r.h.s.) of (2), the field T_n is *stationary*, in the sense that its covariance (2) depends only on the difference $x - y$. From now on, without loss of generality, we assume that T_n is stochastically independent of T_m for $n \neq m$.

For $n \in S$, we will focus on the *zero set* $T_n^{-1}(0) = \{x \in \mathbb{T} : T_n(x) = 0\}$; recall that, according e.g. to [4], with probability one $T_n^{-1}(0)$ consists of the union of a finite number of rectifiable (random) curves, called *nodal lines*, containing a finite set of isolated singular points. In this manuscript, we are more specifically interested in the following two *local* functionals associated with the nodal set $T_n^{-1}(0)$:

1. the *Leray (or microcanonical) measure* defined as [15, (1.1)]

$$\mathcal{L}_n := \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \text{meas} \{x \in \mathbb{T} : |T_n(x)| < \varepsilon\}, \quad (3)$$

where ‘meas’ stands for the Lebesgue measure on \mathbb{T} , and the limit is in the sense of convergence in probability;

2. the (total) *nodal length* \mathcal{L}_n , given by (see [11])

$$\mathcal{L}_n := \text{length}(T_n^{-1}(0)); \quad (4)$$

for technical reasons, we will sometimes need to consider *restricted nodal lengths*, that are defined as follows: for every measurable $Q \subset \mathbb{T}$,

$$\mathcal{L}_n(Q) := \text{length}(T_n^{-1}(0) \cap Q). \quad (5)$$

We observe that, in the jargon of stochastic calculus, the quantity \mathcal{Z}_n corresponds to the *occupation density at zero* of T_n – see [9] for a classical reference on the subject.

As already discussed, our aim is to establish quantitative limit theorems for both \mathcal{Z}_n and \mathcal{L}_n in the *high-energy limit*, that is, when $\mathcal{N}_n \rightarrow +\infty$.

Notation. Given two positive sequences $\{a_n\}_{n \in S}$, $\{b_n\}_{n \in S}$ we will write:

1. $a_n \ll b_n$, if there exists a finite constant $C > 0$ such that $a_n \leq Cb_n$, $\forall n \in S$. Similarly, $a_n \ll_{\alpha} b_n$ (resp. $a_n \ll_{\alpha, \beta} b_n$) will mean that C depends on α (resp. α, β);
2. “ $a_n \ll b_n$, as $\mathcal{N}_n \rightarrow +\infty$ ” (or equivalently “ $a_n = O(b_n)$, as $\mathcal{N}_n \rightarrow +\infty$ ”) if, for every subsequence $\{n\} \subset S$ such that $\mathcal{N}_n \rightarrow \infty$, the ratio a_n/b_n is asymptotically bounded. Similarly, “ $a_n \ll_{\alpha} b_n$, as $\mathcal{N}_n \rightarrow +\infty$ ”, (resp. “ $a_n \ll_{\alpha, \beta} b_n$, as $\mathcal{N}_n \rightarrow +\infty$ ”) will mean that the bounding constant depends on α (resp. α, β);
3. $a_n \asymp b_n$ (resp. $a_n \asymp b_n$, $\mathcal{N}_n \rightarrow +\infty$) if both $a_n \ll b_n$ and $b_n \ll a_n$ (resp. $a_n \ll b_n$ and $b_n \ll a_n$, as $\mathcal{N}_n \rightarrow +\infty$) hold;
4. $a_n = o(b_n)$ if $a_n/b_n \rightarrow 0$ as $n \rightarrow +\infty$ (and analogously for subsequences);
5. $a_n \sim b_n$ if $a_n/b_n \rightarrow 1$ as $n \rightarrow +\infty$ (and analogously for subsequences).

1.2 Previous work

1.2.1 Leray measure

The Leray measure in (3) was investigated by Oravecz, Rudnick and Wigman [15]. They found that [15, Theorem 4.1], for every $n \in S$,

$$\mathbb{E}[\mathcal{Z}_n] = \frac{1}{\sqrt{2\pi}}, \quad (6)$$

i.e. the expected Leray measure is constant, and moreover [15, Theorem 1.1],

$$\text{Var}(\mathcal{Z}_n) = \frac{1}{4\pi\mathcal{N}_n} + O\left(\frac{1}{\mathcal{N}_n^2}\right). \quad (7)$$

In particular, the asymptotic behaviour of the variance, as $\mathcal{N}_n \rightarrow +\infty$, is independent of the distribution of lattice points lying on the circle of radius \sqrt{n} .

1.2.2 Nodal length

The expected nodal length was computed in [16] to be, for $n \in S$,

$$\mathbb{E}[\mathcal{L}_n] = \frac{1}{2\sqrt{2}}\sqrt{E_n}. \quad (8)$$

Computing the nodal variance is a subtler issue, and its asymptotic behaviour (in the high-energy limit) was fully characterized in [11] as follows. We start by observing that the set Λ_n induces a probability measure μ_n on the unit circle \mathbb{S}^1 , given by $\mu_n := \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} \delta_{\lambda/\sqrt{n}}$, where δ_{θ} denotes the Dirac mass at $\theta \in \mathbb{S}^1$. One crucial fact is that,

although there exists a density-1 subsequence $\{n_j\} \subset S$ such that $\mu_{n_j} \Rightarrow d\theta/2\pi$, as $j \rightarrow +\infty$ ¹, there is an infinity of other weak-* adherent points for the sequence $\{\mu_n\}_{n \in S}$ — see [12] for a partial classification. In particular, for every $\eta \in [-1, 1]$, there exists a subsequence $\{n_j\} \subset S$ (see [11, 12]) such that

$$\widehat{\mu_{n_j}}(4) \rightarrow \eta, \quad \text{as } j \rightarrow +\infty, \quad (9)$$

where, for a probability measure μ on the unit circle, the symbol $\widehat{\mu}(4)$ stands for the fourth Fourier coefficient $\widehat{\mu}(4) := \int_{\mathbb{S}^1} \theta^{-4} d\mu(\theta)$. Krishnapur, Kurlberg and Wigman in [11] found that, as $\mathcal{N}_n \rightarrow +\infty$,

$$\text{Var}(\mathcal{L}_n) = c_n \frac{E_n}{\mathcal{N}_n^2} (1 + o(1)), \quad (10)$$

where $c_n := (1 + \widehat{\mu_n}(4)^2)/512$. Such a result is in stark contrast with (7): indeed, it shows that the asymptotic variance of the nodal length multiplicatively depends on the distribution of lattice points lying on the circle of radius \sqrt{n} , via the fluctuations of the squared Fourier coefficient $\widehat{\mu_n}(4)^2$; this also entails that the order of magnitude of the variance is E_n/\mathcal{N}_n^2 , since the sequence $\{|\widehat{\mu_n}(4)|\}_n$ is bounded by 1. Plainly, in order to obtain an asymptotic behaviour in (10) that has no multiplicative corrections, one needs to extract a subsequence $\{n_j\} \subset S$ such that $\mathcal{N}_{n_j} \rightarrow +\infty$ and $|\widehat{\mu_{n_j}}(4)|$ converges to some $\eta \in [0, 1]$; in this case, one deduces that $\text{Var}(\mathcal{L}_{n_j}) \sim c(\eta) E_{n_j}/\mathcal{N}_{n_j}^2$, where $c(\eta) := (1 + \eta^2)/512$. Note that if $\mu_{n_j} \Rightarrow \mu$, then $\widehat{\mu_{n_j}}(4) \rightarrow \widehat{\mu}(4)$. By (9), the possible values of the constant $c(\eta)$ span therefore the whole interval $[1/512, 1/128]$.

The second order behavior of the nodal length was investigated in [13]. Let us define, for $\eta \in [0, 1]$, the random variable

$$\mathcal{M}_\eta := \frac{1}{2\sqrt{1+\eta^2}} (2 - (1+\eta)X_1^2 - (1-\eta)X_2^2), \quad (11)$$

where X_1, X_2 are i.i.d. standard Gaussians. Note that \mathcal{M}_η is invariant in law under the transformation $\eta \mapsto -\eta$, so that if $\eta \in [-1, 0]$ we define $\mathcal{M}_\eta := \mathcal{M}_{-\eta}$.

Theorem 1.1 in [13] states that for $\{n_j\} \subset S$ such that $\mathcal{N}_{n_j} \rightarrow +\infty$ and $|\widehat{\mu_{n_j}}(4)| \rightarrow \eta$, as $j \rightarrow +\infty$, one has that

$$\widetilde{\mathcal{L}_{n_j}} \xrightarrow{d} \mathcal{M}_\eta, \quad (12)$$

where \xrightarrow{d} denotes convergence in distribution and, for $n \in S$,

$$\widetilde{\mathcal{L}_n} := \frac{\mathcal{L}_n - \mathbb{E}[\mathcal{L}_n]}{\sqrt{\text{Var}(\mathcal{L}_n)}} \quad (13)$$

¹ From now on, \Rightarrow denotes weak-* convergence of measures and $d\theta$ the uniform measure on \mathbb{S}^1

is the normalized nodal length. Note that (12) is a non-universal and non central limit theorem: indeed, for $\eta \neq \eta'$ the (non Gaussian) laws of the random variables \mathcal{M}_η and $\mathcal{M}_{\eta'}$ in (11) have different supports.

1.3 Main results

The main purpose of this paper is to prove quantitative limit theorems for local functionals of nodal sets of arithmetic random waves, such as the Leray measure in (3) and the nodal length in (4). We will work with the 1-Wasserstein distance (see e.g. [14, §C] and the references therein). Given two random variables X, Y whose laws are μ_X and μ_Y , respectively, the Wasserstein distance between μ_X and μ_Y , written $d_W(X, Y)$, is defined as

$$d_W(X, Y) := \inf_{(A, B)} \mathbb{E}[|A - B|],$$

where the infimum runs over all pairs of random variables (A, B) with marginal laws μ_X and μ_Y , respectively. We will mainly use the dual representation

$$d_W(X, Y) = \sup_{h \in \mathcal{H}} |\mathbb{E}[h(X) - h(Y)]|, \quad (14)$$

where \mathcal{H} denotes the class of Lipschitz functions $h : \mathbb{R} \rightarrow \mathbb{R}$ whose Lipschitz constant is less or equal than 1. Relation (14) implies in particular that, if $d_W(X_n, X) \rightarrow 0$, then $X_n \xrightarrow{d} X$ (the converse implication is false in general). Our first result is a *uniform* bound for the Wasserstein distance between the normalized Leray measure

$$\widetilde{\mathcal{Z}}_n := \frac{\mathcal{Z}_n - \mathbb{E}[\mathcal{Z}_n]}{\sqrt{\text{Var}(\mathcal{Z}_n)}} \quad (15)$$

and a standard Gaussian random variable.

Theorem 1. *We have that, on S ,*

$$d_W(\widetilde{\mathcal{Z}}_n, Z) \ll \mathcal{N}_n^{-1/2}, \quad (16)$$

where $\widetilde{\mathcal{Z}}_n$ is defined in (15), and $Z \sim \mathcal{N}(0, 1)$ is a standard Gaussian random variable. In particular, if $\{n_j\} \subset S$ is such that $\mathcal{N}_{n_j} \rightarrow +\infty$, then $\widetilde{\mathcal{Z}}_{n_j} \xrightarrow{d} Z$.

The following theorem deals with nodal lengths, providing a quantitative counterpart to the convergence result stated in (12).

Theorem 2. *As $\mathcal{N}_n \rightarrow +\infty$, one has that*

$$d_W(\widetilde{\mathcal{Z}}_n, \mathcal{M}_\eta) \ll \mathcal{N}_n^{-1/4} \vee ||\widehat{\mu}_n(4) - \eta|^{1/2}, \quad (17)$$

where $\widetilde{\mathcal{L}}_n$ and \mathcal{M}_η are defined, respectively, in (12) and (11).

Note that (17) entails the limit theorem (12): it is important to observe that, while the arguments exploited in [13] directly used the variance estimates in [11], the proof of (12) provided in the present paper is basically self-contained, except for the use of a highly non-trivial combinatorial estimate by Bombieri and Bourgain [3], appearing in our proof of Lemma 2 below — see Section 5. We also notice that the bound (16) for the Leray measure is uniform on S , whereas the bound (17) for the nodal length holds asymptotically, and depends on the angular distribution of lattice points lying on the circle of radius \sqrt{n} .

By combining the arguments used in the proofs of Theorem 1 and Theorem 2 with the content of [13, Section 4.2], one can also deduce the following multidimensional limit theorem, yielding in particular a form of *asymptotic dependence* between Leray measures and nodal lengths.

Corollary 1. *Let $\{n_j\} \subset S$ be such that $\mathcal{N}_{n_j} \rightarrow +\infty$ and $|\widehat{\mu}_{n_j}(4)| \rightarrow \eta \in [0, 1]$, then*

$$\left(\widetilde{\mathcal{L}}_{n_j}, \widetilde{\mathcal{L}}_{n_j} \right) \xrightarrow{d} \left(Z_1, \frac{q(Z)}{\sqrt{1+\eta^2}} \right),$$

where $Z = Z(\eta) = (Z_1, Z_2, Z_3, Z_4)$ is a centered Gaussian vector with covariance matrix

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{3+\eta}{8} & \frac{1-\eta}{8} & 0 \\ \frac{1}{2} & \frac{1-\eta}{8} & \frac{3+\eta}{8} & 0 \\ 0 & 0 & 0 & \frac{1-\eta}{8} \end{pmatrix},$$

and q is the polynomial $q(z_1, z_2, z_3, z_4) := 1 + z_1^2 - 2z_2^2 - 2z_3^2 - 4z_4^2$.

The details of the proof are left to the reader.

2 Outline of our approach

2.1 About the proofs of the main results

In order to prove Theorem 1 and Theorem 2, we pervasively use *chaotic expansion* techniques (see §3). Since both \mathcal{Z}_n in (3) and \mathcal{L}_n in (4) are finite-variance functionals of a Gaussian field, they can be written as a series, converging in $L^2(\mathbb{P})$, whose terms can be explicitly found:

$$\mathcal{Z}_n = \sum_{q=0}^{+\infty} \mathcal{Z}_n[2q], \quad \mathcal{L}_n = \sum_{q=0}^{+\infty} \mathcal{L}_n[2q]. \quad (18)$$

For each $q \geq 0$, the random variable $\mathcal{Z}_n[2q]$ (resp. $\mathcal{L}_n[2q]$) is the orthogonal projection of \mathcal{Z}_n (resp. \mathcal{L}_n) onto the so-called *Wiener chaos* of order $2q$, that will be denoted by C_{2q} . Since $C_0 = \mathbb{R}$, we have $\mathcal{Z}_n[0] = \mathbb{E}[\mathcal{Z}_n]$ and $\mathcal{L}_n[0] = \mathbb{E}[\mathcal{L}_n]$; moreover, chaoses of different orders are orthogonal in $L^2(\mathbb{P})$.

2.1.1 On the proof of Theorem 1

We first need the following result, that will be proved in §4.

Proposition 1. *For $n \in S$ (cf. (7))*

$$\text{Var}(\mathcal{Z}_n[2]) = \frac{1}{4\pi\mathcal{N}_n}. \quad (19)$$

Moreover, for every $K \geq 2$,

$$\sum_{q \geq K} \text{Var}(\mathcal{Z}_n[2q]) \ll_K \int_{\mathbb{T}} r_n(x)^{2K} dx \quad \text{on } S; \quad (20)$$

in particular, for $K = 2$,

$$\sum_{q \geq 2} \text{Var}(\mathcal{Z}_n[2q]) \ll \mathcal{N}_n^{-2}. \quad (21)$$

Proposition 1 gives an alternative proof of (7) via chaotic expansions and entails also that, as $\mathcal{N}_n \rightarrow +\infty$,

$$\frac{\mathcal{Z}_n - \mathbb{E}[\mathcal{Z}_n]}{\sqrt{\text{Var}(\mathcal{Z}_n)}} = \frac{\mathcal{Z}_n[2]}{\sqrt{\text{Var}(\mathcal{Z}_n[2])}} + o_{\mathbb{P}}(1),$$

where $o_{\mathbb{P}}(1)$ denotes a sequence converging to 0 in probability. In particular, the Leray measure and its second chaotic component have the same asymptotic behavior, since different order Wiener chaoses are orthogonal. Let us now introduce some more notation. If \sqrt{n} is an integer, we define

$$\Lambda_n^+ := \{\lambda = (\lambda_1, \lambda_2) \in \Lambda_n : \lambda_2 > 0\} \cup \{(\sqrt{n}, 0)\},$$

otherwise $\Lambda_n^+ := \{\lambda = (\lambda_1, \lambda_2) \in \Lambda_n : \lambda_2 > 0\}$. Note that $|\Lambda_n^+| = \mathcal{N}_n/2$ in both cases.

Lemma 1. *For $n \in S$*

$$\mathcal{Z}_n[2] = -\frac{1}{\sqrt{2\pi}} \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n^+} (|a_\lambda|^2 - 1).$$

Lemma 1, proven in §4 below, states that the second chaotic component is (proportional to) a sum of independent random variables. To conclude the proof of Theorem 1, note that we can write

$$d_W(\widetilde{\mathcal{L}}_n, Z) \leq d_W(\widetilde{\mathcal{L}}_n, \widetilde{\mathcal{L}}_n[2]) + d_W(\widetilde{\mathcal{L}}_n[2], Z), \quad (22)$$

where $\widetilde{\mathcal{L}}_n[2] := \mathcal{L}_n[2]/\sqrt{\text{Var}(\mathcal{L}_n[2])}$. The first term on the right-hand side of (22) may be bounded by (21), whereas for the second term standard results apply, thanks to Lemma 1.

2.1.2 On the proof of Theorem 2

The proof of Theorem 2 is similar to that one of Theorem 1. In [13] it has been shown that $\mathcal{L}_n[2] = 0$ for every $n \in S$, and moreover that, as $\mathcal{N}_n \rightarrow +\infty$,

$$\text{Var}(\mathcal{L}_n) \sim \text{Var}(\mathcal{L}_n[4]), \quad (23)$$

by proving that the asymptotic variance of $\mathcal{L}_n[4]$ equals the r.h.s. of (10). The result stated in (23) and the orthogonality properties of Wiener chaoses entail that the fourth chaotic component and the total length have the same asymptotic behavior i.e., as $\mathcal{N}_n \rightarrow +\infty$,

$$\frac{\mathcal{L}_n - \mathbb{E}[\mathcal{L}_n]}{\sqrt{\text{Var}(\mathcal{L}_n)}} = \frac{\mathcal{L}_n[4]}{\sqrt{\text{Var}(\mathcal{L}_n[4])}} + o_{\mathbb{P}}(1), \quad (24)$$

where $o_{\mathbb{P}}(1)$ denotes a sequence converging to 0 in probability. Finally, in [13] it was shown that $\mathcal{L}_n[4]$ can be written as a *polynomial transform* of an asymptotically Gaussian random vector, so that the same convergence as in (12) holds when replacing the total nodal length with its fourth chaotic component.

Now let $h : \mathbb{R} \rightarrow \mathbb{R}$ be a 1-Lipschitz function and $\{n_j\}_j \subset S$ be such that $\mathcal{N}_{n_j} \rightarrow +\infty$ and $|\widehat{\mu}_{n_j}(4)| \rightarrow \eta$, as $j \rightarrow +\infty$. Bearing in mind (14) and (24), we write, by virtue of the triangle inequality,

$$\left| \mathbb{E} \left[h(\widetilde{\mathcal{L}}_{n_j}) - h(\mathcal{M}_\eta) \right] \right| \leq \mathbb{E} \left[\left| h(\widetilde{\mathcal{L}}_{n_j}) - h(\widetilde{\mathcal{L}}_{n_j}[4]) \right| \right] + \mathbb{E} \left[\left| h(\widetilde{\mathcal{L}}_{n_j}[4]) - h(\mathcal{M}_\eta) \right| \right], \quad (25)$$

where $\widetilde{\mathcal{L}}_{n_j}[4] := \mathcal{L}_{n_j}[4]/\sqrt{\text{Var}(\mathcal{L}_{n_j}[4])}$. Let us deal with the first term on the r.h.s. of (25).

Proposition 2. *Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be a 1-Lipschitz function and $\{n\} \subset S$ such that $\mathcal{N}_n \rightarrow +\infty$, then*

$$\mathbb{E} \left[\left| h(\widetilde{\mathcal{L}}_n) - h(\widetilde{\mathcal{L}}_n[4]) \right| \right] \ll \mathcal{N}_n^{-1/4}. \quad (26)$$

In order to prove Proposition 2 in §5, we need to control the behavior of the variance tail $\sum_{q \geq 3} \text{Var}(\mathcal{L}_n[2q])$.

Lemma 2. *For every $K \geq 3$, on S we have*

$$\sum_{q \geq K} \text{Var}(\mathcal{L}_n[2q]) \ll_K E_n \int_{\mathbb{T}} r_n(x)^{2K} dx; \quad (27)$$

in particular, if $\mathcal{N}_n \rightarrow +\infty$,

$$\sum_{q \geq 3} \text{Var}(\mathcal{L}_n[2q]) \ll E_n \mathcal{N}_n^{-5/2}. \quad (28)$$

The proof of Lemma 2 is considerably more delicate than that of (20), see §5, and together with a precise investigation of the fourth chaotic component gives also an alternative proof of (10) via chaotic expansions.

For the second term on the r.h.s. of (25), recall from above that in [13] it was shown that $\mathcal{L}_n[4]$ can be written as a polynomial transform p of a random vector, say $W(n)$, which is asymptotically Gaussian. Let us denote by Z this limiting vector. Then, we can reformulate our problem as the estimation of the distributional distance between $p(W(n_j))$ and $p(Z)$, the latter distributed as \mathcal{M}_η in (11). To prove the following in §6 we can take advantage of some results in [6, 7].

Proposition 3. *Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be a 1-Lipschitz function and let $|\widehat{\mu}_{n_j}(4)| \rightarrow \eta \in [0, 1]$, as $\mathcal{N}_{n_j} \rightarrow +\infty$, then*

$$\left| \mathbb{E} \left[h(\widetilde{\mathcal{L}}_{n_j}[4]) - h(\mathcal{M}_\eta) \right] \right| \ll \mathcal{N}_{n_j}^{-1/4} \vee ||\widehat{\mu}_{n_j}(4)| - \eta|^{1/2}. \quad (29)$$

Proposition 2 and Proposition 3 allow one to prove Theorem 2 in §6, bearing in mind (14) and (25). We now state and prove a technical result, which is a key tool for the proofs of Theorems 1 and 2.

2.2 A technical result

Some of the main bounds in our paper will follow from technical estimates involving pairs of cubes contained in the Cartesian product $\mathbb{T} \times \mathbb{T}$, that will be implicitly classified (for every fixed $n \in S$) according to the behaviour of the mapping $(x, y) \mapsto \mathbb{E}[T_n(x) \cdot T_n(y)] = r_n(x - y)$ appearing in (2).

Notation. For every integer $M \geq 1$, we denote by $\mathcal{Q}(M)$ the partition of \mathbb{T} obtained by translating in the directions k/M ($k \in \mathbb{Z}^2$) the square $Q_0 = Q_0(M) := [0, 1/M) \times [0, 1/M)$. Note that, by construction, $|\mathcal{Q}(M)| = M^2$.

Now we fix, for the rest of the paper, a small number $\varepsilon \in (0, 10^{-3})$. The following statement unifies several estimates taken from [5, §6.1] (yielding Point 4), and [11, §4.1] (yielding Point 5) and [15, §6.1]. A sketch of the proof is provided for the sake of completeness.

Proposition 4. *There exists a mapping $M : S \rightarrow \mathbb{N} : n \mapsto M(n)$, as well as sets $G_0(n), G_1(n) \subset \mathcal{Q}(M(n)) \times \mathcal{Q}(M(n))$ with the following properties:*

1. *there exist constants $1 < c_1 < c_2 < \infty$ such that $c_1 E_n \leq M(n)^2 \leq c_2 E_n$ for every $n \in S$;*
2. *for every $n \in S$, $G_0(n) \cap G_1(n) = \emptyset$ and $G_0(n) \cup G_1(n) = \mathcal{Q}(M(n)) \times \mathcal{Q}(M(n))$;*

3. $(Q, Q') \in G_0(n)$ if and only if for every $(x, y) \in Q \times Q'$, and for every choice of $i \in \{1, 2\}$ and $(i, j) \in \{1, 2\}^2$,

$$|r_n(x - y)|, |\partial_i r_n(x - y)/\sqrt{n}|, |\partial_{i,j} r_n(x - y)/n| \leq \varepsilon, \quad (30)$$

where $\partial_i r_n := \partial/\partial x_i r_n$ and $\partial_{i,j} := \partial/\partial x_i \partial x_j r_n$.

4. for every fixed $K \geq 2$, one has that

$$|G_1(n)| \ll_{\varepsilon, K} E_n^2 \int_{\mathbb{T}} |r_n(x)|^{2K} dx; \quad (31)$$

5. adopting the notation (5), one has that

$$\text{Var}(\mathcal{L}_n(Q_0)) \ll 1/E_n; \quad (32)$$

6. for every fixed $q \geq 2$, one has that

$$\int_{\hat{Q}_0} r_n(x)^{2q} dx \ll \frac{1}{2E_n(q+1)} \left(1 - \left(1 - \frac{E_n}{M(n)^2} \right)^{q+1} \right), \quad (33)$$

where $\hat{Q}_0 := Q_0 - Q_0$, and the constant involved in the above estimates is independent of q .

Proof (Sketch). The combination of Points 1–4 in the above statement corresponds to a slight variation of [5, Lemma 6.3]. Both estimates (32) and (33) follow from the fact that \hat{Q}_0 is contained in the union of four adjacent *positive singular cubes*, in the sense of [15, Definition 6.3]². Using such a representation of \hat{Q}_0 , in order to prove (32) it is indeed sufficient to apply the same arguments as in [11, §4.1] for deducing that, defining the rescaled correlation 2-points function K_2 as in [11, formula (29)],

$$\text{Var}(\mathcal{L}_n(Q_0)) = E_n \int_{Q_0} \int_{Q_0} K_2(x - y) dx dy \leq \frac{E_n}{M(n)^2} \int_{\hat{Q}_0} K_2(x) dx \ll \frac{1}{E_n}.$$

Finally, arguing as in [15, §6.5], we infer that $r_n(x)^2 \leq 1 - E_n \|x - x_0\|^2$, where $x_0 = (0, 0)$ and the estimate holds for every $x \in \hat{Q}_0$, yielding in turn the relations

$$\begin{aligned} \int_{\hat{Q}_0} r_n(x)^{2q} dx &\leq \int_{\|x - x_0\| \leq \frac{1}{M}} (1 - E_n \|x - x_0\|^2)^q dx \ll \int_0^{\frac{1}{M}} r (1 - E_n r^2)^q dr \\ &= \frac{1}{2E_n} \frac{1}{q+1} \left(1 - \left(1 - \frac{E_n}{M(n)^2} \right)^{q+1} \right), \end{aligned}$$

and therefore the desired conclusion. \square

² Indeed, each one of the four cubes composing \hat{Q}_0 is such that its boundary contains the point $x_0 = (0, 0)$, and the singularity in the sense of [15, Definition 6.3] follows by the continuity of trigonometric functions.

3 Local functionals and Wiener chaos

As mentioned in §2.1, for the proof of our main results we need the notion of Wiener-Itô chaotic expansions for non-linear functionals of Gaussian fields. In what follows, we will present it in a simplified form adapted to our situation; we refer the reader to [14, §2.2] for a complete discussion.

3.1 Wiener Chaos

Let ϕ denote the standard Gaussian density on \mathbb{R} and $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \phi(t)dt) =: L^2(\phi)$ the space of square integrable functions on the real line w.r.t. the Gaussian measure $\phi(t)dt$. The sequence of normalized Hermite polynomials $\{(k!)^{-1/2}H_k\}_{k \geq 0}$ is a complete orthonormal basis of $L^2(\phi)$; recall [14, Definition 1.4.1] that they are defined recursively as follows: $H_0 \equiv 1$, and, for $k \geq 1$, $H_k(t) = tH_{k-1}(t) - H'_{k-1}(t)$, $t \in \mathbb{R}$. Recall now the definition of the arithmetic random waves (1), and observe that it involves a family of complex-valued Gaussian random variables $\{a_\lambda : \lambda \in \mathbb{Z}^2\}$ with the following properties: (i) $a_\lambda = b_\lambda + ic_\lambda$, where b_λ and c_λ are two independent real-valued centered Gaussian random variables with variance $1/2$; (ii) a_λ and $a_{\lambda'}$ are independent whenever $\lambda' \notin \{\lambda, -\lambda\}$, and (iii) $a_\lambda = \overline{a_{-\lambda}}$. Consider now the space of all real finite linear combinations of random variables ξ of the form $\xi = za_\lambda + \bar{z}a_{-\lambda}$, where $\lambda \in \mathbb{Z}^2$ and $z \in \mathbb{C}$. Let us denote by \mathbf{A} its closure in $L^2(\mathbb{P})$; it turns out that \mathbf{A} is a real centered Gaussian Hilbert subspace of $L^2(\mathbb{P})$.

Definition 1. Let q be a nonnegative integer; the q -th *Wiener chaos* associated with \mathbf{A} , denoted by C_q , is the closure in $L^2(\mathbb{P})$ of all real finite linear combinations of random variables of the form

$$H_{p_1}(\xi_1) \cdot H_{p_2}(\xi_2) \cdots H_{p_k}(\xi_k)$$

for $k \geq 1$, where the integers $p_1, \dots, p_k \geq 0$ satisfy $p_1 + \dots + p_k = q$, and (ξ_1, \dots, ξ_k) is a standard real Gaussian vector extracted from \mathbf{A} (note that, in particular, $C_0 = \mathbb{R}$).

It is well-known (see [14, §2.2]) that C_q and C_m are orthogonal in $L^2(\mathbb{P})$ whenever $q \neq m$, and moreover $L^2(\Omega, \sigma(\mathbf{A}), \mathbb{P}) = \bigoplus_{q \geq 0} C_q$; equivalently, every real-valued functional F of \mathbf{A} can be (uniquely) represented in the form

$$F = \sum_{q=0}^{\infty} F[q], \tag{34}$$

where $F[q]$ is the orthogonal projection of F onto C_q , and the series converges in $L^2(\mathbb{P})$. Plainly, $F[0] = \mathbb{E}[F]$.

3.2 Chaotic expansion of \mathcal{Z}_n

We can rewrite (3) as

$$\mathcal{Z}_n = \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \int_{\mathbb{T}} 1_{[-\varepsilon, \varepsilon]}(T_n(x)) dx =: \lim_{\varepsilon \rightarrow 0} \mathcal{Z}_n^\varepsilon, \quad (35)$$

and hence formally represent the Leray measure as

$$\mathcal{Z}_n = \int_{\mathbb{T}} \delta_0(T_n(x)) dx, \quad (36)$$

where δ_0 denotes the Dirac mass at $0 \in \mathbb{R}$. Let us now consider the sequence of coefficients $\{\beta_{2q}\}_{q \geq 0}$ defined as

$$\beta_{2q} := \frac{1}{\sqrt{2\pi}} H_{2q}(0), \quad (37)$$

where H_{2q} denotes the $2q$ -th Hermite polynomial, as before. It can be seen as the sequence of coefficients corresponding to the (formal) chaotic expansion of the Dirac mass.

The following result concerns the chaotic expansion of the Leray measure in (36) and will be proved in the Appendix.

Lemma 3. *For $n \in S$, one has that $\mathcal{Z}_n \in L^2(\mathbb{P})$, and the chaotic expansion of \mathcal{Z}_n is*

$$\mathcal{Z}_n = \sum_{q=0}^{+\infty} \mathcal{Z}_n[2q] = \sum_{q=0}^{+\infty} \frac{\beta_{2q}}{(2q)!} \int_{\mathbb{T}} H_{2q}(T_n(x)) dx, \quad (38)$$

where β_{2q} is given in (37), and the convergence of the above series holds in $L^2(\mathbb{P})$.

3.3 Chaotic expansion of \mathcal{L}_n

We recall now from [13] the chaotic expansion (34) for the nodal length. First, \mathcal{L}_n in (4) admits the following integral representation

$$\mathcal{L}_n = \int_{\mathbb{T}} \delta_0(T_n(x)) |\nabla T_n(x)| dx, \quad (39)$$

where δ_0 still denotes the Dirac mass at $0 \in \mathbb{R}$ and ∇T_n the gradient of T_n ; more precisely, $\nabla T_n = (\partial_1 T_n, \partial_2 T_n)$ with $\partial_i := \partial/\partial x_i$ for $i = 1, 2$. The integral in (39) has to be interpreted in the sense that, for any sequence of bounded probability densities $\{g_k\}$ such that the associated probabilities weakly converge to δ_0 , one has that $\int_{\mathbb{T}} g_k(T_n(x)) |\nabla T_n(x)| dx \rightarrow \mathcal{L}_n$ in $L^2(\mathbb{P})$. A straightforward differentiation of the definition (1) of T_n yields, for $j = 1, 2$

$$\partial_j T_n(x) = \frac{2\pi i}{\sqrt{\mathcal{N}_n}} \sum_{(\lambda_1, \lambda_2) \in \Lambda_n} \lambda_j a_{\lambda} e_{\lambda}(x). \quad (40)$$

Hence the random fields $T_n, \partial_1 T_n, \partial_2 T_n$ viewed as collections of Gaussian random variables indexed by $x \in \mathbb{T}$ are all lying in \mathbf{A} , i.e. for every $x \in \mathbb{T}$ we have

$$T_n(x), \partial_1 T_n(x), \partial_2 T_n(x) \in \mathbf{A}.$$

It has been proved in [11] that the random variables $T_n(x), \partial_1 T_n(x), \partial_2 T_n(x)$ are independent for fixed $x \in \mathbb{T}$, and for $i = 1, 2$

$$\text{Var}(\partial_i T_n(x)) = \frac{E_n}{2}. \quad (41)$$

We can write from (39), keeping in mind (41),

$$\mathcal{L}_n = \sqrt{\frac{E_n}{2}} \int_{\mathbb{T}} \delta_0(T_n(x)) |\tilde{\nabla} T_n(x)| dx, \quad (42)$$

with $\tilde{\nabla} T_n := (\tilde{\partial}_1 T_n, \tilde{\partial}_2 T_n)$ and for $i = 1, 2$, $\tilde{\partial}_i := \partial_i / \sqrt{E_n/2}$. Note that $\tilde{\partial}_i T_n(x)$ has unit variance for every $x \in \mathbb{T}$.

Equation (39), or equivalently (42), explicitly represents the nodal length as a (finite-variance) non-linear functional of a Gaussian field. To recall its chaotic expansion, we need (37) and moreover have to introduce the collection of coefficients $\{\alpha_{2n, 2m} : n, m \geq 1\}$, that is related to the Hermite expansion of the norm $|\cdot|$ in \mathbb{R}^2 :

$$\alpha_{2n, 2m} = \sqrt{\frac{\pi}{2}} \frac{(2n)!(2m)!}{n!m!} \frac{1}{2^{n+m}} p_{n+m} \left(\frac{1}{4} \right), \quad (43)$$

where for $N = 0, 1, 2, \dots$ and $x \in \mathbb{R}$

$$p_N(x) := \sum_{j=0}^N (-1)^j \cdot (-1)^N \binom{N}{j} \frac{(2j+1)!}{(j!)^2} x^j,$$

$\frac{(2j+1)!}{(j!)^2}$ being the so-called “swinging factorial” restricted to odd indices. From [13, Proposition 3.2], we have for $q = 2$ or $q = 2m + 1$ odd ($m \geq 1$) $\mathcal{L}_n[q] \equiv 0$, and for $q \geq 2$

$$\begin{aligned} \mathcal{L}_n[2q] &= \sqrt{\frac{4\pi^2 n}{2}} \sum_{u=0}^q \sum_{k=0}^u \frac{\alpha_{2k, 2u-2k} \beta_{2q-2u}}{(2k)!(2u-2k)!(2q-2u)!} \times \\ &\times \int_{\mathbb{T}} H_{2q-2u}(T_n(x)) H_{2k}(\tilde{\partial}_1 T_n(x)) H_{2u-2k}(\tilde{\partial}_2 T_n(x)) dx. \end{aligned} \quad (44)$$

The Wiener-Itô chaotic expansion of \mathcal{L}_n is hence

$$\begin{aligned} \mathcal{L}_n = \mathbb{E}[\mathcal{L}_n] + \sqrt{\frac{4\pi^2 n}{2}} \sum_{q=2}^{+\infty} \sum_{u=0}^q \sum_{k=0}^u \frac{\alpha_{2k,2u-2k} \beta_{2q-2u}}{(2k)!(2u-2k)!(2q-2u)!} \times \\ \times \int_{\mathbb{T}} H_{2q-2u}(T_n(x)) H_{2k}(\tilde{\partial}_1 T_n(x)) H_{2u-2k}(\tilde{\partial}_2 T_n(x)) dx, \end{aligned}$$

with convergence in $L^2(\mathbb{P})$.

3.3.1 Fourth chaotic components

In this part we investigate the fourth chaotic component $\mathcal{L}_n[4]$ (from (44) with $q = 2$), recalling also some facts from [13].

Consider, for $n \in S$, the four-dimensional random vector $W = W(n)$ given by

$$W(n) = \begin{pmatrix} W_1(n) \\ W_2(n) \\ W_3(n) \\ W_4(n) \end{pmatrix} := \frac{1}{\sqrt{\mathcal{N}_n/2}} \sum_{\lambda \in \Lambda_n^+} (|a_\lambda|^2 - 1) \begin{pmatrix} 1 \\ \lambda_1^2/n \\ \lambda_2^2/n \\ \lambda_1 \lambda_2/n \end{pmatrix},$$

whose covariance matrix is

$$\Sigma_n = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{3+\widehat{\mu}_n(4)}{8} & \frac{1-\widehat{\mu}_n(4)}{8} & 0 \\ \frac{1}{2} & \frac{1-\widehat{\mu}_n(4)}{8} & \frac{3+\widehat{\mu}_n(4)}{8} & 0 \\ 0 & 0 & 0 & \frac{1-\widehat{\mu}_n(4)}{8} \end{pmatrix}, \quad (45)$$

see [13, Lemma 4.1]. Note that for every $n \in S$

$$W_2(n) + W_3(n) = W_1(n). \quad (46)$$

The following will be proved in the Appendix and is a finer version of [13, Lemma 4.2].

Lemma 4. *For every $n \in S$,*

$$\mathcal{L}_n[4] = \sqrt{\frac{E_n}{\mathcal{N}_n^2}} \frac{1}{\sqrt{512}} \left(W_1^2 - 2W_2^2 - 2W_3^2 - 4W_4^2 + \frac{1}{2} \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} |a_\lambda|^4 \right), \quad (47)$$

and moreover,

$$\text{Var}(\mathcal{L}_n[4]) = \frac{E_n}{512 \mathcal{N}_n^2} \left(1 + \widehat{\mu}_n(4)^2 + \frac{34}{\mathcal{N}_n} \right). \quad (48)$$

It is worth noticing that Lemma 2 and (48) immediately give an alternative proof of (10) via chaotic expansion.

We recall here from [13, Lemma 4.3] that, for $\{n_j\} \subseteq S$ such that $\mathcal{N}_{n_j} \rightarrow +\infty$ and $\hat{\mu}_{n_j}(4) \rightarrow \eta \in [-1, 1]$, as $j \rightarrow \infty$, the following CLT holds:

$$W(n_j) \xrightarrow{d} Z = Z(\eta) = \begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \end{pmatrix}, \quad (49)$$

where $Z(\eta)$ is a centered Gaussian vector with covariance

$$\Sigma = \Sigma(\eta) = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{3+\eta}{8} & \frac{1-\eta}{8} & 0 \\ \frac{1}{2} & \frac{1-\eta}{8} & \frac{3+\eta}{8} & 0 \\ 0 & 0 & 0 & \frac{1-\eta}{8} \end{pmatrix}. \quad (50)$$

The eigenvalues of Σ are $0, \frac{3}{2}, \frac{1-\eta}{8}, \frac{1+\eta}{4}$ and hence, in particular, Σ is singular. Moreover,

$$\frac{\mathcal{L}_{n_j}[4]}{\sqrt{\text{Var}(\mathcal{L}_{n_j}[4])}} \xrightarrow{d} \mathcal{M}_{|\eta|},$$

where $\mathcal{M}_{|\eta|}$ is defined as in (11), see [13, Proposition 2.2].

4 Proof of Theorem 1

Note first that, from (37) and (38) for $q = 0$

$$\mathcal{Z}_n[0] = \beta_0 = \frac{1}{\sqrt{2\pi}},$$

cf. (6). Let us now focus on the second chaotic component of the Leray measure in (38), by proving Lemma 1.

Proof (Lemma 1). By (37) and (38) for $q = 1$, recalling that $H_2(t) = t^2 - 1$,

$$\mathcal{Z}_n[2] = -\frac{1}{2\sqrt{2\pi}} \int_{\mathbb{T}} (T_n(x)^2 - 1) dx.$$

Finally, (1) allows us to conclude the proof. \square

We can now prove Proposition 1.

Proof (Proposition 1). From Lemma 1, straightforward computations based on independence yield that

$$\text{Var}(\mathcal{Z}_n[2]) = \frac{1}{4\pi\mathcal{N}_n},$$

that is (19). We can rewrite (20) as

$$\sum_{q \geq K} \text{Var}(\mathcal{Z}_n[2q]) = \sum_{q=K}^{+\infty} \frac{\beta_{2q}^2}{(2q)!} \int_{\mathbb{T}} r_n(x)^{2q} dx \ll \int_{\mathbb{T}} r_n(x)^{2K} dx \quad (51)$$

(note that the first equality in (51) is a direct consequence of (38), [14, Proposition 1.4.2] and stationarity of T_n). Our proof of the second equality in (51), which is (20), uses the content of Proposition 4. We can rewrite the middle term in (51), by stationarity of T_n , as

$$\begin{aligned} \sum_{q=K}^{+\infty} \frac{\beta_{2q}^2}{(2q)!} \int_{\mathbb{T}} r_n(x)^{2q} dx &= \sum_{q=K}^{+\infty} \frac{\beta_{2q}^2}{(2q)!} \int_{\mathbb{T}} \int_{\mathbb{T}} r_n(x-y)^{2q} dx dy \\ &= \sum_{q=K}^{+\infty} \frac{\beta_{2q}^2}{(2q)!} \sum_{(Q, Q') \in G_0(n)} \int_Q \int_{Q'} r_n(x-y)^{2q} dx dy \\ &\quad + \sum_{q=K}^{+\infty} \frac{\beta_{2q}^2}{(2q)!} \sum_{(Q, Q') \in G_1(n)} \int_Q \int_{Q'} r_n(x-y)^{2q} dx dy \\ &=: A(n) + B(n). \end{aligned} \quad (52)$$

Using Point 3 in Proposition 4 one infers that

$$\begin{aligned} A(n) &\leq \sum_{q=K}^{+\infty} \frac{\beta_{2q}^2}{(2q)!} \varepsilon^{2q-2K} \sum_{(Q, Q') \in G_0(n)} \int_Q \int_{Q'} r_n(x-y)^{2K} dx dy \\ &\leq \sum_{q=K}^{+\infty} \frac{\beta_{2q}^2}{(2q)!} \varepsilon^{2q-2K} \int_{\mathbb{T}} r_n(x)^{2K} dx. \end{aligned} \quad (53)$$

It is easy to check that, since $\varepsilon \in (0, 1)$, then

$$\sum_{q=1}^{+\infty} \frac{\beta_{2q}^2}{(2q)!} \varepsilon^{2q} < \infty$$

(indeed, $\beta_{2q}^2/(2q)! \asymp 1/\sqrt{q}$, as $q \rightarrow \infty$), finally yielding

$$A(n) \ll_{\varepsilon, K} \int_{\mathbb{T}} r_n(x)^{2K} dx. \quad (54)$$

Let us now focus on $B(n)$. For every pair $(Q, Q') \in G_1(n)$ and every $q \geq 1$, we can use Cauchy-Schwartz inequality and then exploit the stationarity of T_n to write

$$\begin{aligned} \int_Q \int_{Q'} r_n(x-y)^{2q} dx dy &= (2q)!^{-1} \mathbb{E} \left[\int_Q H_{2q}(T_n(x)) dx \int_{Q'} H_{2q}(T_n(y)) dy \right] \\ &\leq (2q)!^{-1} \text{Var} \left(\int_{Q_0} H_{2q}(T_n(x)) dx \right) = \int_{Q_0} \int_{Q_0} r_n(x-y)^{2q} dx dy \\ &\leq \int_{Q_0} dy \int_{\hat{Q}_0} r_n(x)^{2q} dx \ll \frac{1}{E_n} \int_{\hat{Q}_0} r_n(x)^{2q} dx, \end{aligned}$$

where the constant involved in the last estimate is independent of q . Using (31) and (33), one therefore deduces that

$$\begin{aligned} B(n) &\ll \int_{\mathbb{T}} r_n(x)^{2K} dx \times \sum_{q=K}^{+\infty} \frac{\beta_{2q}^2}{(2q)!} \frac{1}{q+1} \left(1 - \left(1 - \frac{E_n}{M^2}\right)^{q+1}\right) \\ &= \int_{\mathbb{T}} r_n(x)^{2K} dx \times \left(\sum_{q=K}^{+\infty} \frac{\beta_{2q}^2}{(2q)!} \frac{1}{q+1} - \sum_{q=K}^{+\infty} \frac{\beta_{2q}^2}{(2q)!} \frac{1}{q+1} \left(1 - \frac{E_n}{M^2}\right)^{q+1} \right). \end{aligned} \quad (55)$$

Since the series appearing in the above expression are both convergent, substituting (53) and (55) in (52), bearing in mind (51), we immediately have (20). To prove (21), it suffices to recall (from (2)) that for every integer $K \geq 1$

$$\int_{\mathbb{T}} r_n(x)^{2K} dx = \frac{|S_{2K}(n)|}{\mathcal{N}_n^{2K}}, \quad (56)$$

where

$$S_{2K}(n) = \{(\lambda_1, \lambda_2, \dots, \lambda_{2K}) \in \Lambda_n^{2K} : \lambda_1 + \lambda_2 + \dots + \lambda_{2K} = 0\}. \quad (57)$$

For $K = 2$, from [11] we have

$$|S_4(n)| = 3\mathcal{N}_n(\mathcal{N}_n - 1), \quad (58)$$

so that substituting (58) into (20) for $K = 2$, bearing in mind (56), we obtain (21). \square

This section ends with the proof of Theorem 1.

Proof (Theorem 1). We write for (22)

$$\begin{aligned} d_W(\widetilde{\mathcal{Z}}_n, Z) &\leq d_W(\widetilde{\mathcal{Z}}_n, \widetilde{\mathcal{Z}}_n[2]) + d_W(\widetilde{\mathcal{Z}}_n[2], Z) \\ &\leq d_W(\widetilde{\mathcal{Z}}_n, \mathcal{Z}_n[2]/\sqrt{\text{Var}(\mathcal{Z}_n)}) + d_W(\mathcal{Z}_n[2]/\sqrt{\text{Var}(\mathcal{Z}_n)}, \widetilde{\mathcal{Z}}_n[2]) \\ &\quad + d_W(\widetilde{\mathcal{Z}}_n[2], Z). \end{aligned} \quad (59)$$

Bearing in mind (14), the first term on the r.h.s. of (59) can be dealt with as follows

$$d_W(\widetilde{\mathcal{Z}}_n, \mathcal{Z}_n[2]/\sqrt{\text{Var}(\mathcal{Z}_n)}) \leq \sqrt{\frac{\sum_{q \geq 2} \text{Var}(\mathcal{Z}_n[2q])}{\text{Var}(\mathcal{Z}_n)}} \ll \mathcal{N}_n^{-1/2}, \quad (60)$$

where the last estimate comes from (21), and the trivial lower bound for the total variance $\text{Var}(\mathcal{Z}_n) \geq \text{Var}(\mathcal{Z}_n[2])$. For the second term on the r.h.s. of (59) we have

$$d_W(\mathcal{Z}_n[2]/\sqrt{\text{Var}(\mathcal{Z}_n)}, \widetilde{\mathcal{Z}}_n[2]) \leq \left| \frac{1}{\sqrt{1 + \frac{\sum_{q \geq 2} \text{Var}(\mathcal{Z}_n[2q])}{\text{Var}(\mathcal{Z}_n[2])}}} - 1 \right| \ll \mathcal{N}_n^{-1}, \quad (61)$$

where we used (19) and (21). Thanks to Lemma 1, we can now deal with the last term in (59) by using the standard Berry-Esseen theorem (see e.g. [14, Section 3.7]). \square

5 Proof of Proposition 2

In this section we will prove Proposition 2. Let us first give the proof of Lemma 2.

Proof (Lemma 2). Fix $K \geq 3$, and recall the notation (5). In order to simplify the discussion, for every $n \in S$ and given $Q \in \mathcal{Q}(M(n))$, we shall denote by $\mathcal{L}_n(Q; \geq 2K)$, the projection of the random variable $\mathcal{L}_n(Q)$ onto the direct sum of chaoses $\bigoplus_{q \geq K} C_{2q}$. For the l.h.s. of (27) we write

$$\sum_{q \geq K} \text{Var}(\mathcal{L}_n[2q]) = \sum_{(Q, Q')} \text{Cov} \{ \mathcal{L}_n(Q; \geq 2K), \mathcal{L}_n(Q'; \geq 2K) \},$$

where the sum runs over the cartesian product $\mathcal{Q}(M(n)) \times \mathcal{Q}(M(n))$. We now write $\sum_{(Q, Q')} = \sum_{(Q, Q') \in G_0(n)} + \sum_{(Q, Q') \in G_1(n)}$, and study separately the two terms. By virtue of Cauchy-Schwarz and stationarity of T_n , one has that

$$\begin{aligned} \sum_{(Q, Q') \in G_1(n)} \text{Cov} \{ \mathcal{L}_n(Q; \geq 2K), \mathcal{L}_n(Q'; \geq 2K) \} &\leq |G_1(n)| \text{Var}(\mathcal{L}_n(Q_0)) \\ &\ll E_n \int_{\mathbb{T}} r_n(x)^{2K} dx, \end{aligned}$$

where we have used (31) and (32), together with the fact that, by orthogonality, $\text{Var}(\mathcal{L}_n(Q; \geq 2K)) \leq \text{Var}(\mathcal{L}_n(Q)) = \text{Var}(\mathcal{L}_n(Q_0))$. The rest of the proof follows closely the arguments rehearsed in [5, §6.2.2]. For all $Q \in \mathcal{Q}(M(n))$, we write

$$\begin{aligned} \mathcal{L}_n(Q; \geq 2K) &= \sqrt{\frac{E_n}{2}} \sum_{q \geq K} \sum_{i_1 + i_2 + i_3 = 2q} \frac{\beta_{i_1} \alpha_{i_2, i_3}}{i_1! i_2! i_3!} \times \\ &\quad \times \int_Q H_{i_1}(T_n(x)) H_{i_2}(\tilde{\partial}_1 T_n(x)) H_{i_3}(\tilde{\partial}_2 T_n(x)) dx, \end{aligned}$$

where the sum runs over all even integers $i_1, i_2, i_3 \geq 0$. We have

$$\begin{aligned}
& \left| \sum_{(Q,Q') \in G_0(n)} \text{Cov}(\mathcal{L}_n(Q; \geq 2K), \mathcal{L}_n(Q'; \geq 2K)) \right| \\
& \leq E_n \sum_{q \geq 2K} \sum_{i_1+i_2+i_3=2q} \sum_{a_1+a_2+a_3=2q} \left| \frac{\beta_{i_1} \alpha_{i_2, i_3}}{i_1! i_2! i_3!} \right| \cdot \left| \frac{\beta_{a_1} \alpha_{a_2, a_3}}{a_1! a_2! a_3!} \right| \\
& \quad \times \left| \sum_{(Q,Q') \in G_0(n)} \int_Q \int_{Q'} \mathbb{E} \left[H_{i_1}(T_n(x)) H_{i_2}(\tilde{\partial}_1 T_n(x)) H_{i_3}(\tilde{\partial}_2 T_n(x)) \right. \right. \\
& \quad \left. \left. \times H_{a_1}(T_n(y)) H_{a_2}(\tilde{\partial}_1 T_n(y)) H_{a_3}(\tilde{\partial}_2 T_n(y)) \right] dx dy \right|.
\end{aligned} \tag{62}$$

For $n \in S$, we now introduce the notation

$$(X_0(x), X_1(x), X_2(x)) := (T_n(x), \tilde{\partial}_1 T_n(x), \tilde{\partial}_2 T_n(x)), \quad x \in \mathbb{T}.$$

Applying the Leonov-Shyriaev formulae for cumulants, in a form analogous to [5, Proposition 2.2], we infer that

$$\begin{aligned}
& \left| \sum_{(Q,Q') \in G_0(n)} \text{Cov}(\mathcal{L}_n(Q; \geq 2K), \mathcal{L}_n(Q'; \geq 2K)) \right| \\
& \leq E_n \sum_{q \geq 2K} \sum_{i_1+i_2+i_3=2q} \sum_{a_1+a_2+a_3=2q} \left| \frac{\beta_{i_1} \alpha_{i_2, i_3}}{i_1! i_2! i_3!} \right| \cdot \left| \frac{\beta_{a_1} \alpha_{a_2, a_3}}{a_1! a_2! a_3!} \right| \\
& \quad \times \mathbf{1}_{\{i_1+i_2+i_3=a_1+a_2+a_3\}} \left| U(i_1, i_2, i_3; a_1, a_2, a_3) \right|, \\
& := E_n \times Z,
\end{aligned} \tag{63}$$

where each summand $U = U(i_1, i_2, i_3; a_1, a_2, a_3)$ is the sum of at most $(2q)!$ terms of the type

$$u = \sum_{(Q,Q') \in G_0(n)} \int_Q \int_{Q'} \prod_{u=1}^{2q} R_{l_u, k_u}(x, y) dx dy, \tag{64}$$

with $k_u, l_u \in \{0, 1, 2\}$ and, for $l, k = 0, 1, 2$ and $x, y \in \mathbb{T}$, and we set

$$R_{l,k}(x, y) := \mathbb{E}[X_l(x) X_k(y)] = R_{l,k}(x - y),$$

where the last equality (with obvious notation) emphasises the fact that $R_{l,k}(x, y)$ only depends on the difference $x - y$. We will also exploit the following relation, valid for every even integer p :

$$\int_{\mathbb{T}} R_{l,k}(x)^p dx \leq \int_{\mathbb{T}} r(x)^p dx; \tag{65}$$

also, for $x, y \in \mathbb{T}$, one has $|R_{l,k}(x - y)| \leq 1$, and, for $(x, y) \in Q \times Q'$,

$$|R_{l,k}(x - y)| \leq \varepsilon. \tag{66}$$

Using the properties of $G_0(n)$ put forward in Proposition 4, as well as the fact that the sum defining Z in (64) involves indices $q \geq 2K$, one infers that, for u as in (65),

$$\begin{aligned} |u| &\leq \varepsilon_1^{2q-2K} \sum_{(Q,Q') \in G_0(n)} \int_Q \int_{Q'} \prod_{u=1}^{2K} |R_{l_u, k_u}(x, y)| dx dy \\ &\leq \varepsilon_1^{2q-2K} \int_{\mathbb{T}} \prod_{u=1}^{2K} |R_{l_u, k_u}(x)| dx \leq \varepsilon_1^{2q-2K} R_n(2K), \end{aligned}$$

where $R_n(2K) = \int_{\mathbb{T}} r_n(x)^{2K} dx$, and we have applied a generalised Hölder inequality together with (66) in order to obtain the last estimate. This relation yields that each of the terms U contributing to Z can be bounded as follows:

$$\begin{aligned} &\left| U(i_1, i_2, i_3; a_1, a_2, a_3) \right| \\ &\leq (2q)! \frac{R_n(2K)}{\varepsilon^{2K}} \varepsilon^{2q} = (2q)! \frac{R_n(2K)}{\varepsilon^{2K}} (\sqrt{\varepsilon})^{i_1+i_2+i_3} (\sqrt{\varepsilon})^{a_1+a_2+a_3}. \end{aligned}$$

This yields that

$$\begin{aligned} Z &\leq \frac{R_n(2K)}{\varepsilon^{2K}} \sum_{q \geq 2K} (2q)! \sum_{i_1+i_2+i_3=2q} \sum_{a_1+a_2+a_3=2q} \left| \frac{\beta_{i_1} \alpha_{i_2, i_3}}{i_1! i_2! i_3!} \right| \times \\ &\quad \left| \frac{\beta_{a_1} \alpha_{a_2, a_3}}{a_1! a_2! a_3!} \right| \times (\sqrt{\varepsilon})^{i_1+i_2+i_3} (\sqrt{\varepsilon})^{a_1+a_2+a_3} := \frac{R_n(2K)}{\varepsilon^{2K}} \times S. \end{aligned}$$

The fact that $S < \infty$ now follows from standard estimates, such as the ones appearing in [5, end of §6.2.2]. This concludes the proof of (27). To prove (28), it suffices to recall (56) for $K = 3$, and use an estimate by Bombieri-Bourgain (see [3, Theorem 1]), stating that $|S_6(n)| = O(\mathcal{N}_n^{7/2})$, as $\mathcal{N}_n \rightarrow +\infty$. \square

We are now ready to prove Proposition 2.

Proof (Proposition 2). By the triangle inequality, for the l.h.s. of (26) we write

$$\begin{aligned} \mathbb{E} \left[\left| h(\widetilde{\mathcal{L}}_n) - h(\widetilde{\mathcal{L}}_n[4]) \right| \right] &\leq \mathbb{E} \left[\left| h(\widetilde{\mathcal{L}}_n) - h(\mathcal{L}_n[4]/\sqrt{\text{Var}(\mathcal{L}_n)}) \right| \right] \\ &\quad + \mathbb{E} \left[\left| h(\mathcal{L}_n[4]/\sqrt{\text{Var}(\mathcal{L}_n)}) - h(\widetilde{\mathcal{L}}_n[4]) \right| \right]. \end{aligned} \quad (68)$$

For the first term on the r.h.s. of (68), since h is Lipschitz, from (18) and Cauchy-Schwartz

$$\begin{aligned} \mathbb{E} \left[\left| h(\widetilde{\mathcal{L}}_n) - h(\mathcal{L}_n[4]/\sqrt{\text{Var}(\mathcal{L}_n)}) \right| \right] &\leq \frac{1}{\sqrt{\text{Var}(\mathcal{L}_n)}} \mathbb{E} \left[\left| \sum_{q \geq 3} \mathcal{L}_n[2q] \right| \right] \\ &\leq \sqrt{\frac{\sum_{q \geq 3} \text{Var}(\mathcal{L}_n[2q])}{\text{Var}(\mathcal{L}_n)}} \ll \mathcal{N}_n^{-1/4}, \end{aligned}$$

where the last upper bound follows from (10) and Lemma 2. For the second term on the r.h.s. of (68), we have again by the Lipschitz property and some standard steps

$$\begin{aligned}
& \mathbb{E} \left[\left| h(\mathcal{L}_n[4]/\sqrt{\text{Var}(\mathcal{L}_n)}) - h(\widetilde{\mathcal{L}}_n[4]) \right| \right] \\
& \leq \left| \frac{1}{\sqrt{\text{Var}(\mathcal{L}_n)}} - \frac{1}{\sqrt{\text{Var}(\mathcal{L}_n[4])}} \right| \mathbb{E}[|\mathcal{L}_n[4]|] \\
& = \frac{1}{\sqrt{\text{Var}(\mathcal{L}_n[4])}} \left| \frac{1}{\sqrt{1 + \frac{\sum_{q \geq 3} \text{Var}(\mathcal{L}_n[2q])}{\text{Var}(\mathcal{L}_n[4])}}} - 1 \right| \mathbb{E}[|\mathcal{L}_n[4]|] \\
& \leq \left| \frac{1}{\sqrt{1 + \frac{\sum_{q \geq 3} \text{Var}(\mathcal{L}_n[2q])}{\text{Var}(\mathcal{L}_n[4])}}} - 1 \right| \ll \mathcal{N}_n^{-1/4},
\end{aligned}$$

where the last bound comes from (48) and Lemma 2. \square

6 Proofs of Proposition 3 and Theorem 2

Recall (46), then we can rewrite (47) as

$$\mathcal{L}_n[4] = \sqrt{\frac{E_n}{\mathcal{N}_n^2}} \frac{1}{\sqrt{512}} \left(p(\widehat{W}) + \psi_n \right), \quad (69)$$

where

$$\psi_n := \frac{1}{2} \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} (|a_\lambda|^4 - 2), \quad (70)$$

$$\widehat{W} := (W_1, W_2, W_4), \quad (71)$$

and p is the polynomial

$$p(x, y, z) := 1 - x^2 - 4y^2 + 4xy - 4z^2. \quad (72)$$

The following statement is a key step in order to prove Proposition 3.

Lemma 5. *Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be a 1-Lipschitz function, define \widehat{W} as in (71) for a fixed $n \in S$, and select $\eta \in [-1, 1]$. Then, on S ,*

$$\left| \mathbb{E} \left[h \left(p(\widehat{W}) \right) \right] - \mathbb{E} \left[h \left(p(\widehat{Z}) \right) \right] \right| \ll |\widehat{\mu}_n(4) - \eta|^{1/2} \vee \mathcal{N}_n^{-1/4}, \quad (73)$$

where the constant involving in the previous estimation is independent of η and h , p is the second degree polynomial defined in (72) and $\widehat{Z} = \widehat{Z}(\eta) := (Z_1, Z_2, Z_4)$ is defined according to (49).

Proof. We will apply an approximation argument from Ch. Döbler's dissertation [6]. Indeed, according to [6, Proposition 2.7.5, Corollary 2.7.6 and Lemma 2.7.7], to every Lipschitz mapping h as in the statement one can associate a collection of real-valued functions $\{h_\rho : \rho \geq 1\}$, such that the following properties are verified for every ρ : (i) h_ρ equals the convolution of h with a centered Gaussian density with variance $1/\rho^2$, (ii) h_ρ is continuously infinitely differentiable, and $\|h_\rho^{(m)}\|_\infty \leq \rho^{m-1}$ (with $h_\rho^{(m)}$ denoting the m th derivative of h_ρ), and (iii) for every integrable random variable X , one has that $|\mathbb{E}[h(X) - h_\rho(X)]| \leq \rho^{-1}$. From Point (iii) it follows in particular that

$$\left| \mathbb{E} \left[h \left(p(\widehat{W}) \right) \right] - \mathbb{E} \left[h \left(p(\widehat{Z}) \right) \right] \right| \leq \frac{2}{\rho} + \left| \mathbb{E} \left[F_\rho(\widehat{W}) \right] - \mathbb{E} \left[F_\rho(\widehat{Z}) \right] \right| =: \frac{2}{\rho} + B(\rho),$$

with $F_\rho := h_\rho \circ p$. Note that F_ρ is an infinitely differentiable mapping, whose partial derivatives have at most polynomial growth. This implies that we can directly apply the same interpolation and integration by parts argument one can find in [14, Proof of Theorem 6.1.2], to deduce that

$$\begin{aligned} B(\rho) &\leq \underbrace{\sum_{i,j=1}^3 |\widehat{\Sigma}(i,j) - \widehat{\Sigma}_n(i,j)| \mathbb{E}[|\partial_{i,j}^2 F_\rho(\widehat{W}(n))|]}_{:=I_1} \\ &+ \underbrace{\sum_{i,j=1}^3 \sqrt{\mathbb{E}[|\partial_{i,j}^2 F_\rho(\widehat{W}(n))|^2] \mathbb{E}[|\widehat{\Sigma}_n(i,j) - \langle D\widehat{W}_j(n), -DL^{-1}\widehat{W}_i(n) \rangle|^2]}}_{:=I_2}, \end{aligned}$$

where $\partial_{i,j}^2 := \partial^2 / \partial x_i \partial x_j$, D denotes the Malliavin derivative (see [14, Definition 1.1.8]), L^{-1} the inverse of the infinitesimal generator of the Ornstein-Uhlenbeck semigroup (see [14, §1.3]) and $\langle \cdot, \cdot \rangle$ stands for the inner product of an appropriate real separable Hilbert space \mathcal{H} (whose exact definition is immaterial for the present proof). Standard arguments based on hypercontractivity and Point (ii) discussed above (together with the fact that $\rho \geq 1$) yield that $E[|\partial_{i,j}^2 F_\rho(\widehat{W}(n))|^2]^{1/2} \leq C\rho$, for some absolute constant C . In view of these facts, relations (45) and (50) imply therefore that

$$I_1 \ll |\widehat{\mu}_n(4) - \eta|. \quad (74)$$

To deal with I_2 , we can use the upper bound in [14, formula (6.2.6)], together with the fact that each $\widehat{W}_i(n)$ belongs to the second Wiener chaos; it hence remains to compute the fourth cumulant $k_4(\widehat{W}_i(n)) = \mathbb{E}[\widehat{W}_i(n)^4] - 3\mathbb{E}[\widehat{W}_i(n)^2]^2$ for every i (note that these cumulants are necessarily positive). Standard computations yield that,

$$k_4(W_1(n)) \ll \frac{1}{\mathcal{N}_n}, \quad k_4(W_2(n)) \ll \frac{1}{\mathcal{N}_n} \frac{1}{\mathcal{N}_\lambda} \sum_{\lambda} \frac{\lambda_1^8}{n^4},$$

$$k_4(W_4(n)) \ll \frac{1}{\mathcal{N}_n} \frac{1}{\mathcal{N}_\lambda} \sum_{\lambda} \frac{\lambda_1^4 \lambda_2^4}{n^4},$$

from which we deduce

$$I_2 \ll \sqrt{\frac{1}{\mathcal{N}_n}}. \quad (75)$$

We have therefore proved the existence of an absolute constant C such that

$$\left| \mathbb{E} \left[h \left(p(\widehat{W}) \right) \right] - \mathbb{E} \left[h \left(p(\widehat{Z}) \right) \right] \right| \leq C \left\{ \frac{1}{\rho} + \rho \gamma_n \right\},$$

with $\gamma_n := (2\mathcal{N}_n^{1/2})^{-1} |\widehat{\mu}_n(4) - \eta| \leq 1$. Since the right-hand side of the previous inequality is maximised at the point $\rho = \gamma_n^{-1/2}$, we immediately obtain the desired conclusion. \square

Let us now prove Proposition 3.

Proof (Proposition 3). We can rewrite the l.h.s. of (29) as

$$\left| \mathbb{E} \left[h \left(\frac{p(\widehat{W}) + \psi_{n_j}}{\sqrt{1 + \widehat{\mu}_{n_j}(4)^2 + 34/\mathcal{N}_{n_j}}} \right) - h \left(\frac{p(Z)}{\sqrt{1 + \eta^2}} \right) \right] \right|,$$

where for $n \in S$, ψ_n is given in (70). By the triangle inequality,

$$\begin{aligned} & \mathbb{E} \left[\left| h \left(\frac{p(\widehat{W}) + \psi_{n_j}}{\sqrt{1 + \widehat{\mu}_{n_j}(4)^2 + 34/\mathcal{N}_{n_j}}} \right) - h \left(\frac{p(Z)}{\sqrt{1 + \eta^2}} \right) \right| \right] \\ & \leq \mathbb{E} \left[\left| h \left(\frac{p(\widehat{W}) + \psi_{n_j}}{\sqrt{1 + \widehat{\mu}_{n_j}(4)^2 + 34/\mathcal{N}_{n_j}}} \right) - h \left(\frac{p(\widehat{W})}{\sqrt{1 + \widehat{\mu}_{n_j}(4)^2 + 34/\mathcal{N}_{n_j}}} \right) \right| \right] \\ & \quad + \mathbb{E} \left[\left| h \left(\frac{p(\widehat{W})}{\sqrt{1 + \widehat{\mu}_{n_j}(4)^2 + 34/\mathcal{N}_{n_j}}} \right) - h \left(\frac{p(\widehat{W})}{\sqrt{1 + \eta^2}} \right) \right| \right] \\ & \quad + \left| \mathbb{E} \left[h \left(\frac{p(\widehat{W})}{\sqrt{1 + \eta^2}} \right) - h \left(\frac{p(Z)}{\sqrt{1 + \eta^2}} \right) \right] \right| \\ & =: I_{n_j} + J_{n_j} + K_{n_j}. \quad (76) \end{aligned}$$

For the first term we simply have, since h is Lipschitz,

$$I_{n_j} \ll \text{Var}(\psi_{n_j}) = \frac{10}{\mathcal{N}_{n_j}}, \quad (77)$$

where the last equality is (85). Let us now deal with J_{n_j} . By the Lipschitz property,

$$\begin{aligned}
J_{n_j} &\leq \sqrt{1 + \widehat{\mu}_{n_j}(4)^2} \left| \frac{1}{\sqrt{1 + \widehat{\mu}_{n_j}(4)^2 + 34/\mathcal{N}_{n_j}}} - \frac{1}{\sqrt{1 + \eta^2}} \right| \\
&= \sqrt{\frac{1 + \widehat{\mu}_{n_j}(4)^2}{1 + \eta^2}} \left| \frac{1}{\sqrt{1 + \frac{\widehat{\mu}_{n_j}(4)^2 - \eta^2 + 34/\mathcal{N}_{n_j}}{1 + \eta^2}}} - 1 \right| \\
&\ll |\widehat{\mu}_{n_j}(4)^2 - \eta^2| + 34\mathcal{N}_{n_j}^{-1} \ll ||\widehat{\mu}_{n_j}(4)| - \eta| \vee \mathcal{N}_{n_j}^{-1}. \tag{78}
\end{aligned}$$

Finally, note that Lemma 5 and the equality in law $\mathcal{M}_\eta = \mathcal{M}_{-\eta}$ give

$$K_n \ll ||\widehat{\mu}_{n_j}(4)| - \eta|^{1/2} \vee \mathcal{N}_{n_j}^{-1/4}.$$

Plugging the latter bound, (77) and (78) into (76) we conclude the proof of Proposition 3. \square

6.1 Proof of Theorem 2

Proof (Theorem 2). For every $j \geq 1$, reasoning as in (25),

$$\begin{aligned}
\mathbb{E} \left[\left| h(\widetilde{\mathcal{Z}}_{n_j}) - h(\mathcal{M}_\eta) \right| \right] &\leq \mathbb{E} \left[\left| h(\widetilde{\mathcal{Z}}_{n_j}) - h(\widetilde{\mathcal{Z}}_{n_j}[4]) \right| \right] \\
&\quad + \mathbb{E} \left[\left| h(\widetilde{\mathcal{Z}}_{n_j}[4]) - h(\mathcal{M}_\eta) \right| \right] \\
&\ll \mathcal{N}_{n_j}^{-1/4} \vee ||\widehat{\mu}_{n_j}(4)| - \eta|^{1/2},
\end{aligned}$$

where the last step directly follows from Proposition 2 and Proposition 3. \square

Acknowledgements We thank Ch. Döbler for useful discussions, and in particular for pointing out the relevance of [6]. The research leading to this work has been supported by the grant FIR-MTH-PUL-15STAR (STARS) at the University of Luxembourg.

Appendix

Proof (Lemma 3). From [13, Lemma 3.4], we have that the chaotic expansion of $\mathcal{Z}_n^\varepsilon$ is

$$\mathcal{Z}_n^\varepsilon = \sum_{q=0}^{+\infty} \mathcal{Z}_n^\varepsilon[2q] = \sum_{q=0}^{+\infty} \frac{\beta_{2q}^\varepsilon}{(2q)!} \int_{\mathbb{T}} H_{2q}(T_n(x)) dx, \tag{79}$$

where H_{2q} denotes the $2q$ -th Hermite polynomial, and

$$\beta_0^\varepsilon = \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \phi(t) dt, \quad \beta_{2q}^\varepsilon = -\frac{1}{\varepsilon} \phi(\varepsilon) H_{2q-1}(\varepsilon), \quad q \geq 1, \quad (80)$$

ϕ still denoting the Gaussian density. Taking the limit for ε going to 0 in (80) we obtain the collection of coefficients (37), related to the (formal) Hermite expansion of the Dirac mass δ_0 . Note that

$$\sum_{q=1}^{+\infty} \frac{(\beta_{2q})^2}{(2q)!} \int_{\mathbb{T}} r_n(x)^{2q} dx = \frac{1}{2\pi} \int_{\mathbb{T}} \left(\frac{1}{\sqrt{1-r_n(x)^2}} - 1 \right) dx < +\infty, \quad (81)$$

since the collection $\{(\beta_{2q})^2/(2q)!\}_q$ coincides with the sequence of Taylor coefficients of the function $x \mapsto 1/(2\pi\sqrt{1-x^2})$ around zero; thanks to Lemma 5.3 in [15] we have the finiteness of the integral. Therefore the series

$$\sum_{q=0}^{+\infty} \frac{\beta_{2q}}{(2q)!} \int_{\mathbb{T}} H_{2q}(T_n(x)) dx,$$

is a well-defined random variable in $L^2(\mathbb{P})$, its variance being the series on the l.h.s. of (81). Moreover, from [1, 22.14.16] and (81)

$$\sum_{q=1}^{+\infty} \frac{(\beta_{2q}^\varepsilon - \beta_{2q})^2}{(2q)!} \int_{\mathbb{T}} r_n(x)^{2q} dx \leq 2 \sum_{q=1}^{+\infty} \frac{(\beta_{2q})^2}{(2q)!} \int_{\mathbb{T}} r_n(x)^{2q} dx < +\infty,$$

that implies, by the dominated convergence theorem, $\mathcal{Z}_n^\varepsilon \rightarrow \mathcal{Z}_n$, $\varepsilon \rightarrow 0$, in $L^2(\mathbb{P})$. \square

Proof (Lemma 4). From (44) with $q = 2$

$$\begin{aligned} \mathcal{L}_n[4] = & \frac{\sqrt{E_n}}{128\sqrt{2}} \left(8 \int_{\mathbb{T}} H_4(T_n(x)) dx - \int_{\mathbb{T}} H_4(\tilde{\partial}_1 T_n(x)) dx - \int_{\mathbb{T}} H_4(\tilde{\partial}_2 T_n(x)) dx \right. \\ & - 8 \int_{\mathbb{T}} H_2(T_n(x)) H_2(\tilde{\partial}_1 T_n(x)) dx - 8 \int_{\mathbb{T}} H_2(T_n(x)) H_2(\tilde{\partial}_2 T_n(x)) dx \\ & \left. - 2 \int_{\mathbb{T}} H_2(\tilde{\partial}_1 T_n(x)) H_2(\tilde{\partial}_2 T_n(x)) dx \right). \end{aligned}$$

Lemmas 5.2 and 5.5 in [13] together with some straightforward computations allow one to write, from (82),

$$\begin{aligned} \mathcal{L}_n[4] = & \sqrt{\frac{E_n}{\mathcal{N}_n^2}} \frac{1}{128\sqrt{2}} \left(8W_1^2 - 16W_2^2 - 16W_3^2 - 32W_4^2 \right. \\ & \left. + \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} |a_\lambda|^4 \left(-8 + 12 \left(\left(\frac{\lambda_1}{\sqrt{n}} \right)^2 + \left(\frac{\lambda_2}{\sqrt{n}} \right)^2 \right)^2 \right) \right). \end{aligned}$$

Recalling that $\lambda_1^2 + \lambda_2^2 = n$, we obtain (47). Let us now note that we can write

$$= \frac{1}{\mathcal{N}_n/2} \sum_{\lambda, \lambda' \in \Lambda_n^+} \left(1 - \frac{2}{n^2} (\lambda_1 \lambda'_1 + \lambda_2 \lambda'_2)^2 \right) (|a_\lambda|^2 - 1)(|a_{\lambda'}|^2 - 1). \quad (82)$$

Then it is immediate to compute from (82)

$$\mathbb{E} [W_1^2 - 2W_2^2 - 2W_3^2 - 4W_4^2] = -1. \quad (83)$$

Bearing in mind Lemma 4.1 in [13], still from (82) some straightforward computations lead to

$$\mathbb{E} [(W_1^2 - 2W_2^2 - 2W_3^2 - 4W_4^2)^2] = 2 + \widehat{\mu}_n(4)^2 + \frac{48}{\mathcal{N}_n}. \quad (84)$$

From (83) and (84) hence we find

$$\text{Var}(W_1^2 - 2W_2^2 - 2W_3^2 - 4W_4^2) = 1 + \widehat{\mu}_n(4)^2 + \frac{48}{\mathcal{N}_n}.$$

Recalling that $(\sqrt{2}|a_\lambda|)^2$ is distributed as a chi-square random variable with two degrees of freedom,

$$\text{Var} \left(\frac{1}{2} \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} |a_\lambda|^4 \right) = \frac{10}{\mathcal{N}_n}, \quad (85)$$

and moreover

$$\text{Cov} \left(W_1^2 - 2W_2^2 - 2W_3^2 - 4W_4^2, \frac{1}{2} \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} |a_\lambda|^4 \right) = -\frac{12}{\mathcal{N}_n}.$$

This concludes the proof of Lemma 4. \square

References

1. Abramovitz, M., Stegun, I.-A.: Handbook of mathematical functions with formulas, graphs, and mathematical tables. National Bureau of Standards Applied Mathematics Series **55**, Washington, D.C. (1964).
2. Berry, M.V. (2002). Statistics of nodal lines and points in chaotic quantum billiards: perimeter corrections, fluctuations, curvature. *J. Phys. A*, **35**, 3025-3038.
3. Bombieri E., Bourgain J.: A problem on sums of two squares. *Int. Math. Res. Not.* **11**, 3343–3407 (2015).
4. Cheng, S.Y.: Eigenfunctions and nodal sets. *Comment. Math. Helv.* (1) **51**, 43–55 (1976).
5. Dalmao, F., Nourdin I., Peccati G., Rossi M.: Phase singularities in complex arithmetic random waves. Preprint. ArXiv 1608.05631 (2016).
6. Döbler, C.: New developments in Stein's method with applications. Ph. D. Thesis, Ruhr-Universität Bochum (2012).

7. Döbler, C.: The Stein equation beyond the support with applications. In preparation (2017+).
8. Erdős, P., and Hall, R.-R.: . On the angular distribution of Gaussian integers with fixed norm. *Discrete Math.* (1-3) **200**, 87–94 (1999).
9. Geman, D., Horowitz, J.: Local times for real and random functions. *Duke Math. J.* (4) **43**, 809–828 (1976).
10. Hardy, G. H., Wright, E.-M.: An introduction to the theory of numbers. Fifth edition. The Clarendon Press, Oxford University Press, New York (1979).
11. Krishnapur, M., Kurlberg, P., Wigman, I.: Nodal length fluctuations for arithmetic random waves. *Ann. of Math.* (2) **177**, 699–737 (2013)
12. Kurlberg, P., Wigman, I.: On probability measures arising from lattice points on circles. *Math. Ann.* in press. ArXiv: 1501.01995 (2015)
13. Marinucci, D., Peccati G., Rossi M., Wigman I.: Non-universality of nodal length distribution for arithmetic random waves. *Geom. Funct. Anal.* (3) **26**, 926–960 (2016)
14. Nourdin, I., Peccati, G.: Normal approximations with Malliavin calculus. From Stein’s method to universality. *Cambridge Tracts in Mathematics*, 192. Cambridge University Press, Cambridge (2012)
15. Oravecz, F., Rudnick Z., Wigman I.: The Leray measure of nodal sets for random eigenfunctions on the torus. *Ann. Inst. Fourier (Grenoble)* (1) **58** 299–335 (2008).
16. Rudnick, Z., Wigman, I.: On the volume of nodal sets for eigenfunctions of the Laplacian on the torus. *Ann. Henri Poincaré.* (1) **9**, 109–130 (2008)
17. Wigman, I.: Fluctuations of the nodal length of random spherical harmonics. *Communications in Mathematical Physics*, 298(3), 787–831 (2010)
18. Wigman, I.: On the nodal lines of random and deterministic Laplace eigenfunctions. *Spectral geometry*, 285297, *Proc. Sympos. Pure Math.*, 84, Amer. Math. Soc., Providence, RI (2012).
19. Yau, S.T. : Survey on partial differential equations in differential geometry. *Seminar on Differential Geometry*, *Ann. of Math. Stud.* 102, 3–71. Princeton Univ. Press, Princeton (1982).